# RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.**

Application Serial Number: 09/909,207 B

Source: 1FW16

Date Processed by STIC: 1-13-05

# *ENTERED*

IFWI6

**RAW SEQUENCE LISTING**　　　　　　　DATE: 01/13/2005
**PATENT APPLICATION: US/09/909,207B**　　　TIME: 14:02:56

Input Set : **A:\GC450-US.txt**
Output Set: **N:\CRF4\01132005\I909207B.raw**

**SEQUENCE LISTING**
```
   4 (1) GENERAL INFORMATION:
C--> 6      (i) APPLICANT: ANDREE LAHAYE
   7                       ERIC DE BUYL
   8                       PIERRE LEDOUX
   9                       RENE DETROZ
  11     (ii) TITLE OF INVENTION: Xylanase, microorganisms produced it,
  12                          DNA molecule, processes for preparation of this xylanase
  13                          and uses thereof
  15    (iii) NUMBER OF SEQUENCES: 29
  17     (iv) CORRESPONDENCE ADDRESS:
  18          (A) ADDRESSEE: WILLIAN BRINKS HOFER GILSON & LIONE
  19          (B) STREET: 2000 K St., N.W., Suite 200
  20          (C) CITY: Washington
  21          (D) STATE: D.C.
  22          (E) COUNTRY: U.S.A.
  23          (F) ZIP: 20006
  25      (v) COMPUTER READABLE FORM:
  26          (A) MEDIUM TYPE: Floppy disk
  27          (B) COMPUTER: IBM PC compatible
  28          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  29          (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
  31     (vi) CURRENT APPLICATION DATA:
C--> 32          (A) APPLICATION NUMBER: US/09/909,207B
C--> 33          (B) FILING DATE: 19-Jul-2001
  34          (C) CLASSIFICATION:
  36    (vii) PRIOR APPLICATION DATA:
  37          (A) APPLICATION NUMBER:
  38          (B) FILING DATE:
  40   (viii) ATTORNEY/AGENT INFORMATION:
  41          (A) NAME: Wilhlem F. Gadiano, Esq.
  42          (B) REGISTRATION NUMBER: 37,136
  43          (C) REFERENCE/DOCKET NUMBER: 4121-40
  45     (ix) TELECOMMUNICATION INFORMATION:
  46          (A) TELEPHONE: 202-429-0625
  47          (B) TELEFAX: (202) 293-1850
  48          (C) TELEX: 650 383-5605
  50 (2) INFORMATION FOR SEQ ID NO: 1:
  52      (i) SEQUENCE CHARACTERISTICS:
  53          (A) LENGTH: 663 base pairs
  54          (B) TYPE: nucleic acid
  55          (C) STRANDEDNESS: single
  56          (D) TOPOLOGY: linear
```

```
58      (ii) MOLECULE TYPE: DNA (genomic)
60      (vi) ORIGINAL SOURCE:
61          (B) STRAIN: Bacillus
65      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
67 CAAATCGTCA CCGACAATTC CATTGGCAAC CACGATGGCT ATGATTATGA ATTTTGGAAA      60
69 GATAGCGGTG GCTCTGGGAC AATGATTCTC AATCATGGCG GTACGTTCAG TGCCCAATGG     120
71 AACAATGTTA ACAACATATT ATTCCGTAAA GGTAAAAAAT TCAATGAAAC ACAAACACAC     180
73 CAACAAGTTG GTAACATGTC CATAAACTAC GGAGCCAACT TCCAACCAAA TGGTAATGCG     240
75 TATTTATGCG TCTATGGTTG GACTGTTGAC CCTCTTGTCG AATATTATAT TGTCGACAGT     300
77 TGGGGCAACT GGCGTCCACC AGGAGCAACG CCTAGGGGA CCATCACTGT TGATGGAGGA      360
79 ACATATGATA TCTACGAGAC TCTTAGAGTC AATCAACCCT CCATTAAGGG GATTGCCACA     420
81 TTTAAACAAT ATTGGAGTGT TCGAAGATCG AAACGCACGA GTGGCACGAT TTCTGTCAGC     480
83 AACCACTTTA GAGCGTGGGA AAACTTAGGG ATGAATATGG GGAAAATGTA TGAAGTCGCG     540
85 CTTACTGTAG AAGGCTATCA AAGTAGCGGA AGTGCTAATG TATATAGCAA TACACTAAGA     600
87 ATTAACGGTA ACCCTCTCTC AACTATTAGT AATGACGAGA GCATAACTTT GGATAAAAAC     660
89 AAT                                                                   663
91 (2) INFORMATION FOR SEQ ID NO: 2:
93      (i) SEQUENCE CHARACTERISTICS:
94          (A) LENGTH: 663 base pairs
95          (B) TYPE: nucleic acid
96          (C) STRANDEDNESS: single
97          (D) TOPOLOGY: linear
99      (ii) MOLECULE TYPE: DNA (genomic)
101     (vi) ORIGINAL SOURCE:
102         (B) STRAIN: Bacillus
104     (ix) FEATURE:
105         (A) NAME/KEY: mat_peptide
106         (B) LOCATION:1..663
108     (ix) FEATURE:
109         (A) NAME/KEY: CDS
110         (B) LOCATION:1..663
113     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
115 CAA ATC GTC ACC GAC AAT TCC ATT GGC AAC CAC GAT GGC TAT GAT TAT      48
116 Gln Ile Val Thr Asp Asn Ser Ile Gly Asn His Asp Gly Tyr Asp Tyr
117  1               5                  10                  15
119 GAA TTT TGG AAA GAT AGC GGT GGC TCT GGG ACA ATG ATT CTC AAT CAT      96
120 Glu Phe Trp Lys Asp Ser Gly Gly Ser Gly Thr Met Ile Leu Asn His
121             20                  25                  30
123 GGC GGT ACG TTC AGT GCC CAA TGG AAC AAT GTT AAC AAC ATA TTA TTC     144
124 Gly Gly Thr Phe Ser Ala Gln Trp Asn Asn Val Asn Asn Ile Leu Phe
125         35                  40                  45
127 CGT AAA GGT AAA AAA TTC AAT GAA ACA CAA ACA CAC CAA CAA GTT GGT     192
128 Arg Lys Gly Lys Lys Phe Asn Glu Thr Gln Thr His Gln Gln Val Gly
129     50                  55                  60
131 AAC ATG TCC ATA AAC TAC GGA GCC AAC TTC CAA CCA AAT GGT AAT GCG     240
132 Asn Met Ser Ile Asn Tyr Gly Ala Asn Phe Gln Pro Asn Gly Asn Ala
133 65                  70                  75                  80
135 TAT TTA TGC GTC TAT GGT TGG ACT GTT GAC CCT CTT GTC GAA TAT TAT     288
136 Tyr Leu Cys Val Tyr Gly Trp Thr Val Asp Pro Leu Val Glu Tyr Tyr
```

RAW SEQUENCE LISTING                    DATE: 01/13/2005
PATENT APPLICATION:  US/09/909,207B     TIME: 14:02:56

Input Set : A:\GC450-US.txt
Output Set: N:\CRF4\01132005\I909207B.raw

```
137                  85                  90                  95
139 ATT GTC GAC AGT TGG GGC AAC TGG CGT CCA CCA GGA GCA ACG CCT AAG        336
140 Ile Val Asp Ser Trp Gly Asn Trp Arg Pro Pro Gly Ala Thr Pro Lys
141              100                 105                 110
143 GGG ACC ATC ACT GTT GAT GGA GGA ACA TAT GAT ATC TAC GAG ACT CTT        384
144 Gly Thr Ile Thr Val Asp Gly Gly Thr Tyr Asp Ile Tyr Glu Thr Leu
145              115                 120                 125
147 AGA GTC AAT CAA CCC TCC ATT AAG GGG ATT GCC ACA TTT AAA CAA TAT        432
148 Arg Val Asn Gln Pro Ser Ile Lys Gly Ile Ala Thr Phe Lys Gln Tyr
149      130                 135                 140
151 TGG AGT GTT CGA AGA TCG AAA CGC ACG AGT GGC ACG ATT TCT GTC AGC        480
152 Trp Ser Val Arg Arg Ser Lys Arg Thr Ser Gly Thr Ile Ser Val Ser
153 145                 150                 155                 160
155 AAC CAC TTT AGA GCG TGG GAA AAC TTA GGG ATG AAT ATG GGG AAA ATG        528
156 Asn His Phe Arg Ala Trp Glu Asn Leu Gly Met Asn Met Gly Lys Met
157              165                 170                 175
159 TAT GAA GTC GCG CTT ACT GTA GAA GGC TAT CAA AGT AGC GGA AGT GCT        576
160 Tyr Glu Val Ala Leu Thr Val Glu Gly Tyr Gln Ser Ser Gly Ser Ala
161              180                 185                 190
163 AAT GTA TAT AGC AAT ACA CTA AGA ATT AAC GGT AAC CCT CTC TCA ACT        624
164 Asn Val Tyr Ser Asn Thr Leu Arg Ile Asn Gly Asn Pro Leu Ser Thr
165      195                 200                 205
167 ATT AGT AAT GAC GAG AGC ATA ACT TTG GAT AAA AAC AAT                    663
168 Ile Ser Asn Asp Glu Ser Ile Thr Leu Asp Lys Asn Asn
169      210                 215                 220
172 (2)  INFORMATION FOR SEQ ID NO: 3:
174       (i)  SEQUENCE CHARACTERISTICS:
175            (A)  LENGTH: 221 amino acids
176            (B)  TYPE: amino acid
177            (D)  TOPOLOGY: linear
179       (ii) MOLECULE TYPE: protein
180       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
182 Gln Ile Val Thr Asp Asn Ser Ile Gly Asn His Asp Gly Tyr Asp Tyr
183   1                 5                 10                  15
185 Glu Phe Trp Lys Asp Ser Gly Gly Ser Gly Thr Met Ile Leu Asn His
186              20                  25                  30
188 Gly Gly Thr Phe Ser Ala Gln Trp Asn Asn Val Asn Asn Ile Leu Phe
189              35                  40                  45
191 Arg Lys Gly Lys Lys Phe Asn Glu Thr Gln Thr His Gln Gln Val Gly
192      50                  55                  60
194 Asn Met Ser Ile Asn Tyr Gly Ala Asn Phe Gln Pro Asn Gly Asn Ala
195  65                  70                  75                  80
197 Tyr Leu Cys Val Tyr Gly Trp Thr Val Asp Pro Leu Val Glu Tyr Tyr
198              85                  90                  95
200 Ile Val Asp Ser Trp Gly Asn Trp Arg Pro Pro Gly Ala Thr Pro Lys
201              100                 105                 110
203 Gly Thr Ile Thr Val Asp Gly Gly Thr Tyr Asp Ile Tyr Glu Thr Leu
204              115                 120                 125
206 Arg Val Asn Gln Pro Ser Ile Lys Gly Ile Ala Thr Phe Lys Gln Tyr
```

```
207       130                   135                   140
209 Trp Ser Val Arg Arg Ser Lys Arg Thr Ser Gly Thr Ile Ser Val Ser
210 145                   150                   155                   160
212 Asn His Phe Arg Ala Trp Glu Asn Leu Gly Met Asn Met Gly Lys Met
213                   165                   170                   175
215 Tyr Glu Val Ala Leu Thr Val Glu Gly Tyr Gln Ser Ser Gly Ser Ala
216             180                   185                   190
218 Asn Val Tyr Ser Asn Thr Leu Arg Ile Asn Gly Asn Pro Leu Ser Thr
219         195                   200                   205
221 Ile Ser Asn Asp Glu Ser Ile Thr Leu Asp Lys Asn Asn
222     210                   215                   220
224 (2)  INFORMATION FOR SEQ ID NO: 4:
226        (i)  SEQUENCE CHARACTERISTICS:
227             (A)  LENGTH: 744 base pairs
228             (B)  TYPE: nucleic acid
229             (C)  STRANDEDNESS: single
230             (D)  TOPOLOGY: linear
232        (ii)  MOLECULE TYPE: DNA (genomic)
234        (vi)  ORIGINAL SOURCE:
235             (B)  STRAIN: Bacillus
239        (xi)  SEQUENCE DESCRIPTION: SEQ ID NO: 4:
241 ATGAGACAAA AGAAATTGAC GTTGATTTTA GCCTTTTTAG TTTGTTTTGC ACTAACCTTA        60
243 CCTGCAGAAA TAATTCAGGC ACAAATCGTC ACCGACAATT CCATTGGCAA CCACGATGGC       120
245 TATGATTATG AATTTTGGAA AGATAGCGGT GGCTCTGGGA CAATGATTCT CAATCATGGC       180
247 GGTACGTTCA GTGCCCAATG GAACAATGTT AACAACATAT TATTCCGTAA AGGTAAAAAA       240
249 TTCAATGAAA CACAAACACA CCAACAAGTT GGTAACATGT CCATAAACTA CGGAGCCAAC       300
251 TTCCAACCAA ATGGTAATGC GTATTTATGC GTCTATGGTT GGACTGTTGA CCCTCTTGTC       360
253 GAATATTATA TTGTCGACAG TTGGGGCAAC TGGCGTCCAC CAGGAGCAAC GCCTAAGGGG       420
255 ACCATCACTG TTGATGGAGG AACATATGAT ATCTACGAGA CTCTTAGAGT CAATCAACCC       480
257 TCCATTAAGG GGATTGCCAC ATTTAAACAA TATTGGAGTG TTCGAAGATC GAAACGCACG       540
259 AGTGGCACGA TTTCTGTCAG CAACCACTTT AGAGCGTGGG AAAACTTAGG GATGAATATG       600
261 GGGAAAATGT ATGAAGTCGC GCTTACTGTA GAAGGCTATC AAAGTAGCGG AAGTGCTAAT       660
263 GTATATAGCA ATACACTAAG AATTAACGGT AACCCTCTCT CAACTATTAG TAATGACGAG       720
265 AGCATAACTT TGGATAAAAA CAAT                                             744
267 (2)  INFORMATION FOR SEQ ID NO: 5:
269        (i)  SEQUENCE CHARACTERISTICS:
270             (A)  LENGTH: 744 base pairs
271             (B)  TYPE: nucleic acid
272             (C)  STRANDEDNESS: single
273             (D)  TOPOLOGY: linear
275        (ii)  MOLECULE TYPE: DNA (genomic)
277        (vi)  ORIGINAL SOURCE:
278             (B)  STRAIN: Bacillus
280        (ix)  FEATURE:
281             (A)  NAME/KEY: CDS
282             (B)  LOCATION:1..744
284        (ix)  FEATURE:
285             (A)  NAME/KEY: mat_peptide
286             (B)  LOCATION:82..744
```

RAW SEQUENCE LISTING                    DATE: 01/13/2005
PATENT APPLICATION:  **US/09/909,207B**        TIME: 14:02:56

Input Set : **A:\GC450-US.txt**
Output Set: **N:\CRF4\01132005\I909207B.raw**

```
288      (ix) FEATURE:
289            (A) NAME/KEY: sig_peptide
290            (B) LOCATION:1..81
293      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
295 ATG AGA CAA AAG AAA TTG ACG TTG ATT TTA GCC TTT TTA GTT TGT TTT        48
296 Met Arg Gln Lys Lys Leu Thr Leu Ile Leu Ala Phe Leu Val Cys Phe
297 -27     -25                     -20                     -15
299 GCA CTA ACC TTA CCT GCA GAA ATA ATT CAG GCA CAA ATC GTC ACC GAC        96
300 Ala Leu Thr Leu Pro Ala Glu Ile Ile Gln Ala Gln Ile Val Thr Asp
301     -10                     -5                      1               5
303 AAT TCC ATT GGC AAC CAC GAT GGC TAT GAT TAT GAA TTT TGG AAA GAT       144
304 Asn Ser Ile Gly Asn His Asp Gly Tyr Asp Tyr Glu Phe Trp Lys Asp
305                 10                  15                  20
307 AGC GGT GGC TCT GGG ACA ATG ATT CTC AAT CAT GGC GGT ACG TTC AGT       192
308 Ser Gly Gly Ser Gly Thr Met Ile Leu Asn His Gly Gly Thr Phe Ser
309                 25                  30                  35
311 GCC CAA TGG AAC AAT GTT AAC AAC ATA TTA TTC CGT AAA GGT AAA AAA       240
312 Ala Gln Trp Asn Asn Val Asn Asn Ile Leu Phe Arg Lys Gly Lys Lys
313             40                  45                  50
315 TTC AAT GAA ACA CAA ACA CAC CAA CAA GTT GGT AAC ATG TCC ATA AAC       288
316 Phe Asn Glu Thr Gln Thr His Gln Gln Val Gly Asn Met Ser Ile Asn
317     55                  60                  65
319 TAC GGA GCC AAC TTC CAA CCA AAT GGT AAT GCG TAT TTA TGC GTC TAT       336
320 Tyr Gly Ala Asn Phe Gln Pro Asn Gly Asn Ala Tyr Leu Cys Val Tyr
321 70                  75                  80                  85
323 GGT TGG ACT GTT GAC CCT CTT GTC GAA TAT TAT ATT GTC GAC AGT TGG       384
324 Gly Trp Thr Val Asp Pro Leu Val Glu Tyr Tyr Ile Val Asp Ser Trp
325                 90                  95                  100
327 GGC AAC TGG CGT CCA CCA GGA GCA ACG CCT AAG GGG ACC ATC ACT GTT       432
328 Gly Asn Trp Arg Pro Pro Gly Ala Thr Pro Lys Gly Thr Ile Thr Val
329             105                 110                 115
331 GAT GGA GGA ACA TAT GAT ATC TAC GAG ACT CTT AGA GTC AAT CAA CCC       480
332 Asp Gly Gly Thr Tyr Asp Ile Tyr Glu Thr Leu Arg Val Asn Gln Pro
333             120                 125                 130
335 TCC ATT AAG GGG ATT GCC ACA TTT AAA CAA TAT TGG AGT GTT CGA AGA       528
336 Ser Ile Lys Gly Ile Ala Thr Phe Lys Gln Tyr Trp Ser Val Arg Arg
337     135                 140                 145
339 TCG AAA CGC ACG AGT GGC ACG ATT TCT GTC AGC AAC CAC TTT AGA GCG       576
340 Ser Lys Arg Thr Ser Gly Thr Ile Ser Val Ser Asn His Phe Arg Ala
341 150                 155                 160                 165
343 TGG GAA AAC TTA GGG ATG AAT ATG GGG AAA ATG TAT GAA GTC GCG CTT       624
344 Trp Glu Asn Leu Gly Met Asn Met Gly Lys Met Tyr Glu Val Ala Leu
345                 170                 175                 180
347 ACT GTA GAA GGC TAT CAA AGT AGC GGA AGT GCT AAT GTA TAT AGC AAT       672
348 Thr Val Glu Gly Tyr Gln Ser Ser Gly Ser Ala Asn Val Tyr Ser Asn
349                 185                 190                 195
351 ACA CTA AGA ATT AAC GGT AAC CCT CTC TCA ACT ATT AGT AAT GAC GAG       720
352 Thr Leu Arg Ile Asn Gly Asn Pro Leu Ser Thr Ile Ser Asn Asp Glu
353             200                 205                 210
```

**VERIFICATION SUMMARY**                          DATE: 01/13/2005
PATENT APPLICATION:  **US/09/909,207B**          TIME: 14:02:57

Input Set : **A:\GC450-US.txt**
Output Set: **N:\CRF4\01132005\I909207B.raw**

L:6 M:220 C: Keyword misspelled or invalid format, [(i) APPLICANT:]
L:32 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:33 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]